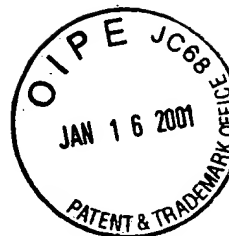


SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: HIGHFIELD, PETER EDMUND
RODGERS, BRIAN COLIN
TEDDER, RICHARD SETON
BARBARA, JOHN ANTHONY JAMES



(ii) TITLE OF INVENTION: VIRAL AGENT

(iii) NUMBER OF SEQUENCES: 25

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: NIXON & VANDERHYE P.C.
(B) STREET: 1100 NORTH GLEBE ROAD
(C) CITY: ARLINGTON
(D) STATE: VIRGINIA
(E) COUNTRY: U.S.A.
(F) ZIP: 22201

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage
(B) COMPUTER: IBM AT Compatible
(C) OPERATING SYSTEM: MS-DOS V3.2
(D) SOFTWARE: Wordperfect 5.0 (DOS text)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 09/664,363
(B) FILING DATE: 18 SEP 2000
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/191,160
(B) FILING DATE: 03 FEB 1994
(C) CLASSIFICATION: 435/235

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/628,516
(B) FILING DATE: 17 DEC 1990
(C) CLASSIFICATION: 435/235

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: UK 89 28 562.1
(B) FILING DATE: 18 DEC 1989

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: UK 90 04 414.0
(B) FILING DATE: 27 FEB 1990

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: UK 90 04 814.1
(B) FILING DATE: 03 MAR 1990

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: MARY J. WILSON
(B) REGISTRATION NUMBER: 32,955
(C) REFERENCE/DOCKET NUMBER: 2035-38

(xi) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (703) 816-4011
(B) TELEFAX: (703) 816-4100

BEST AVAILABLE COPY

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: bacteriophage lambda gt11

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Oligonucleotide synthesizer; oligo d19

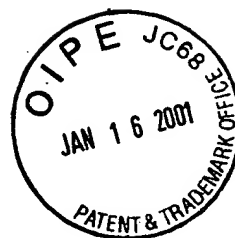
(ix) FEATURE:

- (B) LOCATION: from 1 to 21 bases homologous to upstream portion of lacZ gene flanking the EcoR1 site in bacteriophage lambda gt11
- (D) OTHER INFORMATION: primes DNA synthesis from the phage vector into cDNA inserted at the EcoR1 site.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGTGGCGACG ACTCCTGGAG C

21



(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: bacteriophage lambda gt11

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Oligonucleotide synthesizer; oligo d20

(ix) FEATURE:

- (B) LOCATION: from 1 to 21 bases homologous to downstream portion of lacZ gene flanking the EcoR1 site in bacteriophage lambda gt11
- (D) OTHER INFORMATION: primes DNA synthesis from the phage vector into cDNA inserted at the EcoR1 site.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TTGACACCAG ACCAACTGGT A

21

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1770 base pairs
- (B) TYPE: nucleotide with corresponding protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to genomic RNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: clone JG2 from cDNA library in lambda gt11

(ix) FEATURE:

- (B) LOCATION: from 1 to 1770 bp portion of the PT-NANBH polyprotein
- (D) OTHER INFORMATION: probably encodes viral non-structural proteins

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CAA AAT GAC TTC CCA GAC GCT GAC CTC ATC GAG GCC AAC CTC CTG TGG	48
Gln Asn Asp Phe Pro Asp Ala Asp Leu Ile Glu Ala Asn Leu Leu Trp	
5 10 15	
CGG CAT GAG ATG GGC GGG GAC ATT ACC CGC GTG GAG TCA GAG AAC AAG	96
Arg His Glu Met Gly Gly Asp Ile Thr Arg Val Glu Ser Glu Asn Lys	
20 25 30	
GTA GTA ATC CTG GAC TCT TTC GAC CCG CTC CGA GCG GAG GAG GAT GAG	144
Val Val Ile Leu Asp Ser Phe Asp Pro Leu Arg Ala Glu Glu Asp Glu	
35 40 45	
CGG GAA GTG TCC GTC CCG GCG GAG ATC CTG CGG AAA TCC AAG AAA TTC	192
Arg Glu Val Ser Val Pro Ala Glu Ile Leu Arg Lys Ser Lys Lys Phe	
50 55 60	
CCA CCA GCG ATG CCC GCA TGG GCA CGC CCG GAT TAC AAC CCT CCG CTG	240
Pro Pro Ala Met Pro Ala Trp Ala Arg Pro Asp Tyr Asn Pro Pro Leu	
65 70 75 80	
CTG GAG TCC TGG AAG GCC CCG GAC TAC GTC CCT CCA GTG GTA CAT GGG	288
Leu Glu Ser Trp Lys Ala Pro Asp Tyr Val Pro Pro Val Val His Gly	
85 90 95	

TGC Cys	CCA Pro	CTG Leu	CCA Pro 100	CCT Pro	ACT Thr	ACC Lys	ACC Thr	CCT Pro 105	CCT Pro	ATA Ile	CCA Pro	CCT Pro 110	CGG Arg	AGA Arg	336	
AAG Lys	AGG Arg	ACA Thr 115	GTT Val	GTT Val	CTG Leu	ACA Thr	GAA Glu 120	TCC Ser	ACC Thr	GTG Val	TCT Ser	TCT Ser 125	GCC Ala	CTG Leu	GCG Ala	384
GAG Glu 130	CTT Leu	GCC Ala	ACA Thr	AAG Lys	GCT Ala	TTT Phe 135	GGT Gly	AGC Ser	TCC Ser	GGA Gly	CCG Pro 140	TCG Ser	GCC Ala	GTC Val	GAC Asp	432
AGC Ser 145	GGC Gly	ACG Thr	GCA Ala	ACC Thr	GCC Ala 150	CCT Pro	CCT Pro	GAC Asp	CAA Gln 155	TCC Ser	TCC Ser	GAC Asp	GAC Asp	GGC Gly	GGA Gly 160	480
GCA Ala	GGA Gly	TCT Ser	GAC Asp	GTT Val 165	GAG Glu	TCG Ser	TAT Tyr	TCC Ser 170	TCC Ser	ATG Met	CCC Pro	CCC Pro	CTT Leu	GAG Glu 175	GGG Gly	528
GAG Glu	CCG Pro	GGG Gly	GAC Asp 180	CCC Pro	GAT Asp	CTC Leu	AGC Ser	GAC Asp 185	GGG Gly	TCT Ser	TGG Trp	TCT Ser	ACC Thr 190	GTG Val	AGT Ser	576
GAG Glu	GAG Glu	GCC Ala 195	GGT Gly	GAG Glu	GAC Asp	GTC Val 200	GTC Val	TGC Cys	TGC Cys	TCG Ser	ATG Met	TCC Ser 205	TAC Tyr	ACA Thr	TGG Trp	624
ACA Thr 210	GGC Gly	GCT Ala	CTG Leu	ATC Ile	ACG Thr	CCA Pro 215	TGC Cys	GCT Ala	GCG Ala	GAG Glu	GAA Glu 220	AGC Ser	AAG Lys	CTG Leu	CCC Pro	672
ATC Ile 225	AAC Asn	GCG Ala	TTG Leu	AGC Ser	AAC Asn 230	TCT Ser	TTG Leu	CTG Leu	CGT Arg	CAC His 235	CAC His	AAC Asn	ATG Met	GTC Val	TAC Tyr 240	720
GCT Ala	ACC Thr	ACA Thr	TCC Ser	CGC Arg 245	AGC Ser	GCA Ala	AGC Ser	CAG Gln	CGG Arg 250	CAG Gln	AAG Lys	AAG Lys	GTC Val	ACC Thr 255	TTT Phe	768
GAC Asp	AGA Arg	CTG Leu	CAA Gln 260	ATC Ile	CTG Leu	GAC Asp	GAT Asp	CAC His 265	TAC Tyr	CAG Gln	GAC Asp	GTG Val	CTC Leu 270	AAG Lys	GAG Glu	816
ATG Met	AAG Lys	GCG Ala 275	AAG Lys	GCG Ala	TCC Ser	ACA Thr	GTT Val 280	AAG Lys	GCT Ala	AAG Lys	CTT Leu	CTA Leu 285	TCA Ser	GTA Val	GAG Glu	864
GAA Glu 290	GCC Ala	TGC Cys	AAG Lys	CTG Leu	ACG Thr	CCC Pro 295	CCA Pro	CAT His	TCG Ser	GCC Ala	AAA Lys 300	TCT Ser	AAA Lys	TTT Phe	GGC Gly	912

TAT Tyr 305	GGG Gly	GCA Ala	AAG Lys	GAC Asp	GTC Val 310	CGG Arg	AAC Asn	CTA Leu	TCC Ser	AGC Ser 315	AAG Lys	GCC Ala	ATT Ile	AAC Asn	CAC His 320	960
ATC Ile	CGC Arg	TCC Ser	GTG Val	TGG Trp 325	GAG Glu	GAC Asp	TTG Leu	TTG Leu	GAA Glu 330	GAC Asp	ACT Thr	GAA Glu	ACA Thr	CCA Pro 335	ATT Ile	1008
GAC Asp	ACC Thr	ACC Thr	ATC Ile 340	ATG Met	GCA Ala	AAA Lys	AAT Asn 345	GAG Glu	GTT Val	TTC Phe	TGC Cys	GTC Val	CAA Gln 350	CCA Pro	GAG Glu	1056
AGA Arg	GGA Gly	GGC Gly 355	CGC Arg	AAG Lys	CCA Pro	GCT Ala	CGC Arg 360	CTT Leu	ATC Ile	GTG Val	TTC Phe	CCA Pro 365	GAC Asp	TTG Leu	GGG Gly	1104
GTC Val 370	CGT Arg	GTG Val	TGC Cys	GAG Glu	AAA Lys	ATG Met 375	GCC Ala	CTC Leu	TAT Tyr	GAC Asp 380	GTG Val	GTC Val	TCC Ser	ACC Thr	CTC Leu	1152
CCT Pro 385	CAG Gln	GCT Ala	GTG Val	ATG Met	GGC Gly 390	TCC Ser	TCG Ser	TAC Tyr	GGA Gly	TTC Phe 395	CAG Gln	TAT Tyr	TCT Ser	CCT Pro	GGA Gly 400	1200
CAG Gln	CGG Arg	GTC Val	GAG Glu	TTC Phe 405	CTG Leu	GTG Val	AAC Asn	GCC Ala	TGG Trp 410	AAA Lys	TCA Ser	AAG Lys	AAG Lys	ACC Thr 415	CCT Pro	1248
ATG Met	GGC Gly	TTT Phe	GCA Ala 420	TAT Tyr	GAC Asp	ACC Thr	CGC Arg	TGT Cys 425	TTT Phe	GAC Asp	TCA Ser	ACA Thr	GTC Val 430	ACT Thr	GAG Glu	1296
AAT Asn	GAC Asp	ATC Ile 435	CGT Arg	GTA Val	GAG Glu	GAG Glu	TCA Ser 440	ATT Ile	TAT Tyr	CAA Gln	TGT Cys	TGT Cys	GAC Asp 445	TTG Leu	GCC Ala	1344
CCC Pro 450	GAA Glu	GCC Ala	AGA Arg	CAG Gln	GCC Ala	ATA Ile 455	AGG Arg	TCG Ser	CTC Leu	ACA Thr	GAG Glu 460	CGG Arg	CTT Leu	TAT Tyr	ATC Ile	1392
GGG Gly 465	GGT Gly	CCC Pro	CTG Leu	ACT Thr	AAT Asn 470	TCA Ser	AAA Lys	GGG Gly	CAG Gln	AAC Asn 475	TGC Cys	GGC Gly	TAT Tyr	CGC Arg	CGG Arg 480	1440
TGC Cys	CGC Arg	GCG Ala	AGC Ser	GGC Gly 485	GTG Val	CTG Leu	ACG Thr	ACT Thr	AGC Ser 490	TGC Cys	GGT Gly	AAT Asn	ACC Thr	CTC Leu 495	ACA Thr	1488
TGT Cys	TAC Tyr	TTG Leu	AAG Lys 500	GCC Ala	TCT Ser	GCA Ala	GCC Ala	TGT Cys 505	CGA Arg	GCT Ala	GCA Ala	AAG Lys	CTC Leu 510	CAG Gln	GAC Asp	1536

TGC	ACG	ATG	CTC	GTG	TGC	GAC	GAC	GGC	CTT	GTC	GTT	ATC	TCT	GAG	AGC	1584
Cys	Thr	Met	Leu	Val	Cys	Gly	Asp	Asp	Leu	Val	Val	Ile	Cys	Glu	Ser	
		515					520					525				
GCG	GGA	ACC	CAG	GAG	GAC	GCG	GCG	AGC	CTA	CGA	GTC	TTC	ACG	GAG	GCT	1632
Ala	Gly	Thr	Gln	Glu	Asp	Ala	Ala	Ser	Leu	Arg	Val	Phe	Thr	Glu	Ala	
	530					535					540					
ATG	ACT	AGG	TAC	TCT	GCC	CCC	CCC	GGG	GAC	CCG	CCC	CAA	CCA	GAA	TAC	1680
Met	Thr	Arg	Tyr	Ser	Ala	Pro	Pro	Gly	Asp	Pro	Pro	Gln	Pro	Glu	Tyr	
545					550				555					560		
GAC	CTG	GAG	TTG	ATA	ACA	TCA	TGC	TCC	TCC	AAT	GTG	TCG	GTC	GCG	CAC	1728
Asp	Leu	Glu	Leu	Ile	Thr	Ser	Cys	Ser	Ser	Asn	Val	Ser	Val	Ala	His	
			565					570						575		
GAT	GCA	TCT	GGC	AAA	AGG	GTA	TAC	TAC	CTC	ACC	CGT	GAC	CCG			1770
Asp	Ala	Ser	Gly	Lys	Arg	Val	Tyr	Tyr	Leu	Thr	Arg	Asp	Pro			
			580					585					590			

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1035 base pairs
- (B) TYPE: nucleotide with corresponding protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to genomic RNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: clone JG3 from cDNA library in lambda gt11

(ix) FEATURE:

- (B) LOCATION: from 1 to 1035 bp portion of the PT-NANBH polyprotein
- (D) OTHER INFORMATION: probably encodes viral non-structural proteins

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACA	GAA	GTG	GAT	GGG	GTG	CGG	CTG	CAC	AGG	TAC	GCT	CCG	GCG	TGC	AAA	48
Thr	Glu	Val	Asp	Gly	Val	Arg	Leu	His	Arg	Tyr	Ala	Pro	Ala	Cys	Lys	
				5				10						15		
CCT	CTC	CTA	CGG	GAG	GAG	GTC	ACA	TTC	CAG	GTC	GGG	CTC	AAC	CAA	TAC	96
Pro	Leu	Leu	Arg	Glu	Glu	Val	Thr	Phe	Gln	Val	Gly	Leu	Asn	Gln	Tyr	
			20				25						30			

CTG Leu	GTT Val	GGG Gly 35	TCG Ser	CAG Gln	CTC Leu	CCC Pro	TGC Cys 40	GAG Glu	CCC Pro	GAA Glu	CCG Pro	GAT Asp 45	GCA Val	GCA Ala	GTG Val	144
CTC Leu	ACT Thr 50	TCC Ser	ATG Met	CTC Leu	ACC Thr	GAC Asp 55	CCC Pro	TCC Ser	CAC His	ATC Ile	ACA Thr 60	GCA Ala	GAG Glu	ACG Thr	GCT Ala	192
AAG Lys 65	CGC Arg	AGG Arg	CTG Leu	GCC Ala	AGG Arg 70	GGG Gly	TCT Ser	CCC Pro	CCC Pro	TCC Ser 75	TTG Leu	GCC Ala	AGC Ser	TCT Ser	TCA Ser 80	240
GCT Ala	AGC Ser	CAG Gln	TTG Leu	TCT Ser 85	GGC Gly	CCT Pro	TCC Ser	TCG Ser	AAG Lys 90	GCG Ala	ACA Thr	TAC Tyr	ATT Ile	ACC Thr 95	CAA Gln	288
AAT Asn	GAC Asp	TTC Phe	CCA Pro 100	GAC Asp	GCT Ala	GAC Asp	CTC Leu	ATC Ile 105	GAG Glu	GCC Ala	AAC Asn	CTC Leu	CTG Leu 110	TGG Trp	CGG Arg	336
CAT His	GAG Glu	ATG Met 115	GGC Gly	GGG Gly	GAC Asp	ATT Ile	ACC Thr 120	CGC Arg	GTG Val	GAG Glu	TCA Ser	GAG Glu 125	AAC Asn	AAG Lys	GTA Val	384
GTA Val 130	ATC Ile	CTG Leu	GAC Asp	TCT Ser	TTC Phe	GAC Asp 135	CCG Pro	CTC Leu	CGA Arg	GCG Ala	GAG Glu 140	GAG Glu	GAT Asp	GAG Glu	CGG Arg	432
GAA Glu 145	GTG Val	TCC Ser	GTC Val	CCG Pro	GCG Ala 150	GAG Glu	ATC Ile	CTG Leu	CGG Arg	AAA Lys 155	TCC Ser	AAG Lys	AAA Lys	TTC Phe	CCA Pro 160	480
CCA Pro	GCG Ala	ATG Met	CCC Pro	GCA Ala 165	TGG Trp	GCA Ala	CGC Arg	CCG Pro	GAT Asp 170	TAC Tyr	AAC Asn	CCT Pro	CCG Pro	CTG Leu 175	CTG Leu	528
GAG Glu	TCC Ser	TGG Trp	AAG Lys 180	GCC Ala	CCG Pro	GAC Asp	TAC Tyr	GTC Val 185	CCT Pro	CCA Pro	GTG Val	GTA Val	CAT His 190	GGG Gly	TGC Cys	576
CCA Pro	CTG Leu	CCA Pro 195	CCT Pro	ACT Thr	AAG Lys	ACC Thr	CCT Pro 200	CCT Pro	ATA Ile	CCA Pro	CCT Pro	CCA Pro 205	CGG Arg	AGA Arg	AAG Lys	624
AGG Arg	ACA Thr 210	GTT Val	GTT Val	CTG Leu	ACA Thr	GAA Glu 215	TCC Ser	ACC Thr	GTG Val	TCT Ser	TCT Ser 220	GCC Ala	CTG Leu	GCG Ala	GAG Glu	672
CTT Leu 225	GCC Ala	ACA Thr	AAG Lys	GCT Ala	TTT Phe 230	GGT Gly	AGC Ser	TCC Ser	GGA Gly	CCG Pro 235	TCG Ser	GCC Ala	GTC Val	GAC Asp	AGC Ser 240	720

GGC	ACG	GCA	ACC	GCC	CCT	CCG	GAC	CAA	TCC	TCC	GAC	GAC	GGC	GGA	GCA	768
Gly	Thr	Ala	Thr	Ala	Pro	Pro	Asp	Gln	Ser	Ser	Asp	Asp	Gly	Gly	Ala	
				245					250					255		
GGA	TCT	GAC	GTT	GAG	TCG	TAT	TCC	TCC	ATG	CCC	CCC	CTT	GAG	GGG	GAG	816
Gly	Ser	Asp	Val	Glu	Ser	Tyr	Ser	Ser	Met	Pro	Pro	Leu	Glu	Gly	Glu	
			260					265					270			
CCG	GGG	GAC	CCC	GAT	CTC	AGC	GAC	GGG	TCT	TGG	TCT	ACC	GTG	AGT	GAG	864
Pro	Gly	Asp	Pro	Asp	Leu	Ser	Asp	Gly	Ser	Trp	Ser	Thr	Val	Ser	Glu	
			275				280					285				
GAG	GCC	GGT	GAG	GAC	GTC	GTC	TGC	TGC	TCG	ATG	TCC	TAC	ACA	TGG	ACA	912
Glu	Ala	Gly	Glu	Asp	Val	Val	Cys	Cys	Ser	Met	Ser	Tyr	Thr	Trp	Thr	
	290					295					300					
GGC	GCT	CTG	ATC	ACG	CCA	TGC	GCT	GCG	GAG	GAA	AGC	AAG	CTG	CCC	ATC	960
Gly	Ala	Leu	Ile	Thr	Pro	Cys	Ala	Ala	Glu	Glu	Ser	Lys	Leu	Pro	Ile	
305					310					315					320	
AAC	GCG	TTG	AGC	AAC	TCT	TTG	CTG	CGT	CAC	CAC	AAC	ATG	GTC	TAC	GCT	1008
Asn	Ala	Leu	Ser	Asn	Ser	Leu	Leu	Arg	His	His	Asn	Met	Val	Tyr	Ala	
				325					330					335		
ACC	ACA	TCC	CGC	AGC	GCA	AGC	CAG	CGG								1035
Thr	Thr	Ser	Arg	Ser	Ala	Ser	Gln	Arg								
			340					345								

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 834 base pairs
- (B) TYPE: nucleotide with corresponding protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to genomic RNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: clone BR11 from cDNA library in lambda gt11

(ix) FEATURE:

- (B) LOCATION: from 1 to 834 bp portion of the PT-NANBH polyprotein
- (D) OTHER INFORMATION: probably encodes viral structural proteins

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGA AAA ACC AAA CGT AAC ACC AAC CTC CGC CCA CAG GAC GTC AGG TTC	48
Arg Lys Thr Lys Arg Asn Thr Asn Leu Arg Pro Gln Asp Val Arg Phe	
5 10 15	
CCG GGC GGT GGT CAG ATC GTT GGT GGA GTT TAC CTG TTG CCG CGC AGG	96
Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg	
20 25 30	
GGC CCC AGG TTG GGT GTG CGC GCG ACT AGG AAG ACT TCC GAG CGG TCG	144
Gly Pro Arg Leu Gly Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser	
35 40 45	
CAA CCT CGT GGA AGG CGA CAA CCT ATC CCC AAG GCT CGC CAG CCC GAG	192
Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys Ala Arg Gln Pro Glu	
50 55 60	
GGC AGG GCC TGG GCT CAG CCC GGG TAC CCT TGG CCC CTC TAT GGC AAC	240
Gly Arg Ala Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn	
65 70 75 80	
GAG GGC ATG GGG TGG GCA GGA TGG CTC CTG TCA CCC CGT GGC TCC CGG	288
Glu Gly Met Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg	
85 90 95	
CCT AGT TGG GGC CCC ACT GAC CCC CGG CGT AGG TCG CGT AAT TTG GGT	336
Pro Ser Trp Gly Pro Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly	
100 105 110	
AAA GTC ATC GAT ACC CTC ACA TGC GGC TTC GCC GAC TCT CAT GGG GTA	384
Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Ser His Gly Val	
115 120 125	
CAT TCC GCT CGT CGG CGC TCC CTT AGG GGC GCT GCC AGG GCC CTG GCG	432
His Ser Ala Arg Arg Arg Ser Leu Arg Gly Ala Ala Arg Ala Leu Ala	
130 135 140	
CAT GGC GTC CGG GTT CTG GAG GAC GGC GTG AAC TAT GCA ACA GGG AAT	480
His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala Thr Gly Asn	
145 150 155 160	
TTA CCC GGT TGC TCT TTC TCT ATC TTC CTC TTG GCT TTG CTG TCC TGT	528
Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Leu Ser Cys	
165 170 175	
TTG ACC ATT CCA GCT TCC GCT TAT GAA GTG CGC AAC GTG TCC GGG ATC	576
Leu Thr Ile Pro Ala Ser Ala Tyr Glu Val Arg Asn Val Ser Gly Ile	
180 185 190	
TAC CAT GTC ACG AAC GAT TGC TCC AAC TCA AGC ATC GTG TAC GAG ACA	624
Tyr His Val Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Thr	
195 200 205	

GCG	GAC	ATG	ATC	ATG	CAC	ACC	CCC	GGG	TGT	GTG	CCC	TGT	GTC	CGG	GAG	672
Ala	Asp	Met	Ile	Met	His	Thr	Pro	Gly	Cys	Val	Pro	Cys	Val	Arg	Glu	
210						215					220					
GGT	AAT	TCC	TCC	CGC	TGC	TGG	GTA	GCG	CTC	ACT	CCC	ACG	CTC	GCG	GCC	720
Gly	Asn	Ser	Ser	Arg	Cys	Trp	Val	Ala	Leu	Thr	Pro	Thr	Leu	Ala	Ala	
225					230					235					240	
AAG	GAC	GCC	AGC	ATC	CCC	ACT	GCG	ACA	ATA	CGA	CGC	CAC	GTC	GAT	TTG	768
Lys	Asp	Ala	Ser	Ile	Pro	Thr	Ala	Thr	Ile	Arg	Arg	His	Val	Asp	Leu	
				245					250					255		
CTC	GTT	GGG	GCG	GCT	GCC	TTC	TCG	TCC	GCT	ATG	TAC	GTG	GGG	GAT	CTC	816
Leu	Val	Gly	Ala	Ala	Ala	Phe	Ser	Ser	Ala	Met	Tyr	Val	Gly	Asp	Leu	
			260					265					270			
TGC	GGA	TCT	GTT	TTC	CCG											834
Cys	Gly	Ser	Val	Phe	Pro											
			275													

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: bacteriophage lambda gt11

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Oligonucleotide synthesizer; oligo d75

(ix) FEATURE:

- (B) LOCATION: from 4 to 9 bases BamH1 site, from 10 to 31 bases homologous to upstream portion of lacZ gene flanking the EcoR1 site in bacteriophage lambda gt11 from 26 to 31 bases EcoR1 site
- (D) OTHER INFORMATION: primes DNA synthesis from the phage vector into cDNA inserted at the EcoR1 site and introduces a BamH1 site suitable for subsequent cloning into expression vectors.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TAAGGATCCC CCGTCAGTAT CGGCGGAATT C

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: bacteriophage lambda gt11

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Oligonucleotide synthesizer; oligo d76

(ix) FEATURE:

- (B) LOCATION: from 4 to 9 bases BamH1 site from 10 to 30 bases homologous to downstream portion of lacZ gene flanking the EcoR1 site in bacteriophage lambda gt11
- (D) OTHER INFORMATION: primes DNA synthesis from the phage vector into cDNA inserted at the EcoR1 site and introduces a BamH1 site suitable for subsequent cloning into expression vectors.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TATGGATCCG TAGCGACCGG CGCTCAGCTG

30

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: oligonucleotide synthesizer; oligo d94

(ix) FEATURE:

- (B) LOCATION: from 1 to 19 bases homologous to bases 914 to 932 of the sense strand of JG2 (SEQ ID NO:3)
- (D) OTHER INFORMATION: primes DNA synthesis on the negative strand of PT-NANBH genomic RNA/DNA.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATGGGGCAAA GGACGTCCG

19

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: oligonucleotide synthesizer; oligo d95

(ix) FEATURE:

- (B) LOCATION: from 1 to 24 bases homologous to bases 1620 to 1643 of the anti-sense strand of JG2 (SEQ ID NO:3)
- (D) OTHER INFORMATION: primes DNA synthesis on the positive strand of PT-NANBH genomic RNA/DNA.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TACCTAGTCA TAGCCTCCGT GAAG

24

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: oligonucleotide synthesizer; oligo N1

(ix) FEATURE:

- (B) LOCATION: from 1 to 17 bases homologous to bases 1033 to 1049 of the sense strand of JG2 (SEQ ID NO:3)
- (D) OTHER INFORMATION: primes DNA synthesis on the negative strand of PT-NANBH genomic RNA/DNA.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAGGTTTTCT GCGTCCA

17

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: oligonucleotide synthesizer; oligo N2

(ix) FEATURE:

- (B) LOCATION: from 1 to 17 bases homologous to bases 1421 to 1437 of the anti-sense strand of JG2 (SEQ ID NO:3)
- (D) OTHER INFORMATION: primes DNA synthesis on the positive strand of PT-NANBH genomic RNA/DNA.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCGATAGCCG CAGTTCT

17

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: oligonucleotide synthesizer; oligo d164

(ix) FEATURE:

- (B) LOCATION: from 1 to 22 bases homologous to bases 10 to 31 of the sequence in Fig. 2 of Okamoto et al., Japan. J. Exp. Med., 1990, 60 167-177, base 22 changed from A to T to introduce Bgl2 recognition site from 8 to 13 bases Bgl2 recognition site

- (D) OTHER INFORMATION: primes DNA synthesis on the negative strand of PT-NANBH genomic RNA/DNA and introduces a Bgl2 site.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCACCATAGA TCTCTCCCCT GT

22

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: oligonucleotide synthesizer; oligo d137

(ix) FEATURE:

- (B) LOCATION: from 1 to 30 bases homologous to bases 154 to 183 of the negative strand of BR11 (SEQ ID NO:5) bases 174, 177 and 178 modified to introduce an EcoR1 recognition site from 5 to 10 bases EcoR1 recognition site
- (D) OTHER INFORMATION: primes DNA synthesis on the positive strand of PT-NANBH genomic RNA/DNA and introduces an EcoR1 site for cloning.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GCGAGAATTC GGGATAGGTT GTCGCCTTCC

30

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: oligonucleotide synthesizer; oligo d136

(ix) FEATURE:

(B) LOCATION: from 1 to 27 bases homologous to bases 672 to 698 of the positive strand of BR11 (SEQ ID NO:5) base 675 changed to G to introduce an EcoR1 recognition site from 4 to 9 bases EcoR1 recognition site

(D) OTHER INFORMATION: primes DNA synthesis on the negative strand of PT-NANBH genomic RNA/DNA and introduces an EcoR1 site for cloning.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGGGAATTCC TCCCGCTGCT GGGTAGC

27

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: chimpanzee; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: oligonucleotide synthesizer; oligo d155

(ix) FEATURE:

(B) LOCATION: from 1 to 28 bases homologous to bases 462 to 489 of the negative strand of figure 47, European Patent Application 88310922.5; bases 483 and 485 changed to introduce an EcoR1 recognition site from 5 to 10 bases EcoR1 recognition site

(D) OTHER INFORMATION: primes DNA synthesis on the positive strand of PT-NANBH genomic RNA/DNA and introduces an EcoR1 site for cloning.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ACGGGAATTC GACCAGGCAC CTGGGTGT

28

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: chimpanzee; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: oligonucleotide synthesizer; oligo d156

(ix) FEATURE:

(B) LOCATION: from 1 to 23 bases homologous to bases 3315 to 3337 of the positive strand of figure 47, European Patent Application 88310922.5; base 3323 changed to C to introduce an EcoRI recognition site from 4 to 9 bases EcoRI recognition site

(D) OTHER INFORMATION: primes DNA synthesis on the negative strand of PT-NANBH genomic RNA/DNA and introduces an EcoRI site for cloning.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTTGAATTCT GGGAGGGCGT CTT

23

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 bases

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: oligonucleotide synthesizer; oligo d92

(ix) FEATURE:

(B) LOCATION: from 1 to 29 bases homologous to bases 36 to 64 of the negative strand of JG2 (SEQ ID NO:3); bases 57, 58 and 60 changed to introduce an EcoRI recognition site from 5 to 10 bases EcoRI recognition site

(D) OTHER INFORMATION: primes DNA synthesis on the positive strand of PT-NANBH genomic RNA/DNA and introduces an EcoRI site for cloning.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

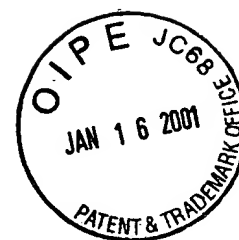
CGCCGAATTC ATGCCGCCAC AGGAGGTTG

29

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 504 base pairs
- (B) TYPE: nucleotide with corresponding protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear



(ii) MOLECULE TYPE: cDNA to genomic RNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: clone 164/137

(ix) FEATURE:

- (B) LOCATION: from 308 to 504 bp start of the PT-NANBH polyprotein
- (D) OTHER INFORMATION: probably encodes viral structural proteins

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GATCACTCCC CTGTGAGGAA CTACTGTCTT CACGCAGAAA GCGTCTAGCC ATGGCGTTAG	60
TATGAGTGTC GTGCAGCCTC CAGGACCCCC CCTCCCGGGA GAGCCATAGT GGTCTGCGGA	120
ACCGGTGAGT ACACCGGAAT TGCCAGGACG ACCGGGTCCT TTCTTGATT AACCCGCTCA	180
ATGCCTGGAG ATTTGGGCGT GCCCCGCAA GACTGCTAGC CGAGTAGTGT TGGGTCGCGA	240
AAGGCCTTGT GGTACTGCCT GATAGGGTGC TTGCGAGTGC CCCGGGAGGT CTCGTAGACC	300
GTGCACC ATG AGC ACG AAT CCT AAA CCT CAA AGA AAA ACC AAA CGT AAC	349
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn	
5 10	
ACC AAC CGC CGC CCA CAG GAC GTC AAG TTC CCG GGC GGT GGT CAG ATC	397
Thr Asn Pro Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile	
15 20 25 30	
GTT GGT GGA GTT TAC CTG TTG CCG CGC AGG GGC CCC AGG TTG GGT GTG	445
Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val	
35 40 45	
CGC GCG ACT AGG AAG ACT TCC GAG CGG TCG CAA CCT CGT GGA AGG CGA	493
Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg	
50 55 60	
CAA CCT ATC CC	504
Gln Pro Ile Pro	
65	

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1107 base pairs
- (B) TYPE: nucleotide with corresponding protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to genomic RNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: clone 136/155

(ix) FEATURE:

- (B) LOCATION: from 1 to 1107 bp portion of the PT-NANBH polyprotein
- (D) OTHER INFORMATION: probably encodes viral structural proteins

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCC TCC CGC TGC TGG GTA GCG CTC ACT CCC ACG CTC GCG GCC AAG GAC	48
Ser Ser Arg Cys Trp Val Ala Leu Thr Pro Thr Leu Ala Ala Lys Asp	
5 10 15	
GCC AGC ATC CCC ACT GCG ACA ATA CGA CGC CAC GTC GAT TTG CTC GTT	96
Ala Ser Ile Pro Thr Ala Thr Ile Arg Arg His Val Asp Leu Leu Val	
20 25 30	
GGG GCG GCT GCC TTC TGC TCC GCT ATG TAC GTG GGG GAT CTC TGC GGA	144
Gly Ala Ala Ala Phe Cys Ser Ala Met Tyr Val Gly Asp Leu Cys Gly	
35 40 45	
TCT GTT TTC CTC GTC TCT CAG CTG TTC ACC TTC TCG CCT CGC CGA CAT	192
Ser Val Phe Leu Val Ser Gln Leu Phe Thr Phe Ser Pro Arg Arg His	
50 55 60	
CAG ACG GTA CAG GAC TGC AAT TGT TCA ATC TAT CCC GGC CAC GTA TCA	240
Gln Thr Val Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Val Ser	
65 70 75 80	
GGT CAC CGC ATG GCT TGG GAT ATG ATG ATG AAC TGG TCA CCT ACA GCA	288
Gly His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Ala	
85 90 95	
GCC CTA GTG GTA TCG CAG CTA CTC CGG ATC CCA CAA GCT GTC GTG GAC	336
Ala Leu Val Val Ser Gln Leu Leu Arg Ile Pro Gln Ala Val Val Asp	
100 105 110	

ATG Met	GTG Val	GCG Ala 115	GGG Gly	GCC Ala	CAC His	TGG Trp	AGA Gly 120	GTC Val	CTG Leu	GCG Ala	GGC Gly	CTT Leu 125	GAC Ala	TAT Tyr	384	
TCC Ser	ATG Met 130	GTG Val	GGG Gly	AAC Asn	TGG Trp	GCT Ala 135	AAG Lys	GTC Val	TTG Leu	GTT Val	GTG Val 140	ATG Met	CTA Leu	CTC Leu	TTT Phe	432
GCC Ala 145	GGC Gly	GTT Val	GAC Asp	GGG Gly	GAA Glu 150	CCT Pro	TAC Tyr	ACG Thr	ACA Thr	GGG Gly 155	GGG Gly	ACA Thr	CAC His	GGC Gly	CGC Arg 160	480
GCC Ala	GCC Ala	CAC His	GGG Gly	CTT Leu 165	ACA Thr	TCC Ser	CTC Leu	TTC Phe	ACA Thr 170	CCT Pro	GGG Gly	CCG Pro	GCT Ala	CAG Gln 175	AAA Lys	528
ATC Ile	CAG Gln	CTT Leu 180	GTA Val	AAC Asn	ACC Thr	AAC Asn	GGC Gly	AGC Ser 185	TGG Trp	CAC His	ATC Ile	AAC Asn	AGA Arg 190	ACT Thr	GCC Ala	576
TTG Leu	AAC Asn 195	TGC Cys	AAT Asn	GAC Asp	TCC Ser	CTC Leu	CAA Gln 200	ACT Thr	GGG Gly	TTC Phe	CTT Leu	GCC Ala 205	GCG Ala	CTG Leu	TTC Phe	624
TAC Tyr 210	ACG Thr	CAC His	AGG Arg	TTC Phe	AAT Asn	GCG Ala 215	TCC Ser	GGA Gly	TGC Cys	TCA Ser	GAG Glu 220	CGC Arg	ATG Met	GCC Ala	AGC Ser	672
TGC Cys 225	CGC Arg	CCC Pro	ATT Ile	GAC Asp	CAG Gln 230	TTC Phe	GAT Asp	CAG Gln	GGG Gly	TGG Trp 235	GGT Gly	CCC Pro	ATC Ile	ACT Thr	TAT Tyr 240	720
AAT Asn	GAG Glu	TCC Ser	CAC His	GGC Gly 245	TTG Leu	GAC Asp	CAG Gln	AGG Arg	CCC Pro 250	TAT Tyr	TGC Cys	TGG Trp	CAC His	TAC Tyr 255	GCA Ala	768
CCT Pro	CAA Gln	CCG Pro	TGT Cys 260	GGT Gly	ATC Ile	GTG Val	CCC Pro	GCG Ala 265	TTG Leu	CAG Gln	GTG Val	TGT Cys	GGC Gly 270	CCA Pro	GTG Val	816
TAC Tyr 275	TGT Cys	TTC Phe	ACT Thr	CCA Pro	AGC Ser	CCT Pro	GTT Val 280	GTG Val	GTG Val	GGG Gly	ACG Thr	ACC Thr	GAT Asp	CGT Arg	TTC Phe	864
GGC Gly 290	GCC Ala	CCT Pro	ACG Thr	TAC Tyr	AGA Arg	TGG Trp 295	GGT Gly	GAG Glu	AAT Asn	GAG Glu	ACG Thr	GAC Asp	GTG Val	CTG Leu	CTT Leu	912
CTC Leu 305	AAC Asn	AAC Asn	ACG Thr	CGG Arg	CCG Pro 310	CCA Pro	CGG Arg	GGC Gly	AAC Asn	TGG Trp 315	TTC Phe	GGC Gly	TGT Cys	ACA Thr	TGG Trp 320	960

ATG AAT AGC ACC GGG TTC ACG AGC TGT GGG GGC CCC GGC AAC	1008
Met Asn Ser Thr Gly Phe Thr Lys Thr Cys Gly Gly Pro Pro Cys Asn	
325 330 335	
ATC GGG GGG GTC GGC AAC AAC ACT TTG ATC TGC CCC ACG GAC TGC TTC	1056
Ile Gly Gly Val Gly Asn Asn Thr Leu Ile Cys Pro Thr Asp Cys Phe	
340 345 350	
CGG AAG CAT CCC GAG GCC ACT TAC ACC AAA TGC GGT TCG GGG CCT TGG	1104
Arg Lys His Pro Glu Ala Thr Tyr Thr Lys Cys Gly Ser Gly Pro Trp	
355 360 365	
TTG	1107
Leu	

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2043 base pairs
- (B) TYPE: nucleotide with corresponding protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to genomic RNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: clone 156/92

(ix) FEATURE:

- (B) LOCATION: from 1 to 2043 bp portion of the PT-NANBH polyprotein
- (D) OTHER INFORMATION: probably encodes viral non-structural proteins

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TGG GAG GGC GTC TTC ACA GGC CTC ACC CAC GTG GAT GCC CAC TTC CTG	48
Trp Glu Gly Val Phe Thr Gly Leu Thr His Val Asp Ala His Phe Leu	
5 10 15	
TCC CAA ACA AAG CAG GCA GGA GAC AAC TTC CCC TAC CTG GTG GCG TAC	96
Ser Gln Thr Lys Gln Ala Gly Asp Asn Phe Pro Tyr Leu Val Ala Tyr	
20 25 30	
CAG GCT ACT GTG TGC GCT AGG GCC CAG GCC CCA CCT CCA TCA TGG GAT	144
Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp	
35 40 45	

CAA Gln 50	ATG Met 50	TGG Trp 50	AAG Lys 50	TGT Cys 50	CTC Leu 50	A Ile 55	EGG Arg 55	CTA Leu 55	AAG Lys 55	CCT Pro 55	ACT Thr 60	CTG Leu 60	GGG Arg 60	CCA Gly 60	192	
ACA Thr 65	CCC Pro 65	TTG Leu 65	CTG Leu 65	TAT Tyr 65	AGG Arg 70	CTG Leu 70	GGA Gly 70	GCC Ala 70	GTC Val 70	CAA Gln 75	AAC Asn 75	GAG Glu 75	GTC Val 75	ACC Thr 80	CTC Leu 80	240
ACA Thr 85	CAC His 85	CCC Pro 85	ATA Ile 85	ACC Thr 85	AAA Lys 85	TTC Phe 85	ATC Ile 85	ATG Met 90	GCA Ala 90	TGC Cys 90	ATG Met 90	TCA Ser 90	GCC Ala 95	GAC Asp 95	CTG Leu 95	288
GAG Glu 100	GTC Val 100	GTC Val 100	ACG Thr 100	AGC Ser 100	ACC Thr 100	TGG Trp 100	GTG Val 105	CTG Leu 105	GTG Val 105	GGC Gly 105	GGG Gly 105	GTC Val 110	CTT Leu 110	GCA Ala 110	GCT Ala 110	336
CTG Leu 115	GCT Ala 115	GCG Ala 115	TAT Tyr 115	TGC Cys 115	TTG Leu 115	ACA Thr 120	ACA Thr 120	GGC Gly 120	AGC Ser 120	GTG Val 120	GTC Val 120	ATT Ile 125	GTG Val 125	GGT Gly 125	AGG Arg 125	384
ATC Ile 130	ATC Ile 130	TTG Leu 130	TCC Ser 130	GGG Gly 130	CGG Arg 130	CCG Pro 135	GCT Ala 135	ATT Ile 135	GTT Val 135	CCC Pro 140	GAC Asp 140	AGG Arg 140	GAA Glu 140	GTC Val 140	CTC Leu 140	432
TAC Tyr 145	CAG Gln 145	GAG Glu 145	TTC Phe 145	GAT Asp 150	GAG Glu 150	ATG Met 150	GAA Glu 150	GAG Glu 150	TGC Cys 155	GCG Ala 155	TCG Ser 155	CAC His 155	CTC Leu 155	CCT Pro 160	TAC Tyr 160	480
ATC Ile 165	GAG Glu 165	CAG Gln 165	GGA Gly 165	ATG Met 165	CAG Gln 165	CTC Leu 165	GCC Ala 170	GAG Glu 170	CAG Gln 170	TTC Phe 170	AAG Lys 170	CAA Gln 170	AAA Lys 170	GCG Ala 175	CTC Leu 175	528
GGG Gly 180	TTG Leu 180	CTG Leu 180	CAG Gln 180	ACA Thr 180	GCC Ala 180	ACC Thr 180	AAG Lys 185	CAA Gln 185	GCG Ala 185	GAG Glu 185	GCC Ala 190	GCT Ala 190	GCT Ala 190	CCC Pro 190	GTG Val 190	576
GTG Val 195	GAG Glu 195	TCC Ser 195	AAG Lys 195	TGG Trp 195	CGA Arg 200	GCC Ala 200	CTT Leu 200	GAG Glu 200	ACC Thr 200	TTC Phe 205	TGG Trp 205	GCG Ala 205	AAA Lys 205	CAC His 205	ATG Met 205	624
TGG Trp 210	AAC Asn 210	TTC Phe 210	ATC Ile 210	AGC Ser 210	GGG Gly 215	ATA Ile 215	CAG Gln 215	TAC Tyr 215	TTA Leu 215	GCA Ala 220	GGC Gly 220	TTG Leu 220	TCC Ser 220	ACT Thr 220	CTG Leu 220	672
CCT Pro 225	GGG Gly 225	AAT Asn 225	CCC Pro 225	GCG Ala 230	ATT Ile 230	GCA Ala 230	TCA Ser 230	CTG Leu 235	ATG Met 235	GCG Ala 235	TTC Phe 235	ACA Thr 235	GCC Ala 240	TCT Ser 240	GTC Val 240	720
ACT Thr 245	AGC Ser 245	CCG Pro 245	CTC Leu 245	ACC Thr 245	ACC Thr 245	CAA Gln 245	TCT Ser 250	ACC Thr 250	CTC Leu 250	CTG Leu 250	CTT Leu 250	AAC Asn 250	ATC Ile 250	CTG Leu 255	GGG Gly 255	768

GGA Gly	TGG Trp	GTA Val	GCC Ala 260	GCC Ala	CAA Gln	CTT Leu	ACT Ala	CCC Pro 265	CCC Pro	AGT Ser	GCT Ala	GCT Ala 270	GCT Ala	TTC Phe	816	
GTA Val	GGC Gly	GCC Ala 275	GGC Gly	ATT Ile	GCT Ala	GGT Gly	GCG Ala 280	GCT Ala	GTT Val	GGC Gly	AGC Ser	ATA Ile 285	GGC Gly	CTT Leu	GGG Gly	864
AAG Lys	GTG Val 290	CTT Leu	GTG Val	GAC Asp	ATC Ile	TTG Leu 295	GCG Ala	GGC Gly	TAT Tyr	GGA Gly 300	GCA Ala	GGA Gly	GTG Val	GCA Ala	GGC Gly	912
GCG Ala 305	CTC Leu	GTG Val	GCC Ala	TTT Phe	AAG Lys 310	GTC Val	ATG Met	AGC Ser	GGC Gly	GAA Glu 315	ATG Met	CCC Pro	TCC Ser	ACC Thr	GAG Glu 320	960
GAC Asp	CTG Leu	GTT Val	AAC Asn	TTA Leu 325	CTC Leu	CCT Pro	GCC Ala	ATC Ile	CTC Leu 330	TCT Ser	CCT Pro	GGT Gly	GCC Ala	CTG Leu 335	GTC Val	1008
GTC Val	GGG Gly	GTC Val	GTG Val 340	TGC Cys	GCA Ala	GCG Ala	ATA Ile	CTG Leu 345	CGT Arg	CGG Arg	CAC His	GTG Val	GGT Gly 350	CCA Pro	GGG Gly	1056
GAG Glu	GGG Gly	GCT Ala 355	GTG Val	CAG Gln	TGG Trp	ATG Met	AAC Asn 360	CGG Arg	CTG Leu	ATA Ile	GCG Ala	TTC Phe 365	GCC Ala	TCG Ser	CGG Arg	1104
GGT Gly 370	AAC Asn	CAT His	GTT Val	TCC Ser	CCC Pro	ACG Thr 375	CAC His	TAT Tyr	GTG Val	CCA Pro	GAG Glu 380	AGC Ser	GAC Asp	GCC Ala	GCA Ala	1152
GCA Ala 385	CGT Arg	GTC Val	ACT Thr	CAG Gln	ATC Ile 390	CTC Leu	TCC Ser	GAC Asp	CTT Leu	ACT Thr 395	ATC Ile	ACC Thr	CAA Gln	CTG Leu	TTG Leu 400	1200
AAG Lys	AGG Arg	CTC Leu	CAC His	CAG Gln 405	TGG Trp	ATT Ile	AAC Asn	GAG Glu 410	GAC Asp	TGC Cys	TCC Ser	ACG Thr	CCC Pro	TGC Cys 415	TCC Ser	1248
GGC Gly	TCG Ser	TGG Trp	CTA Leu 420	AGG Arg	GAT Asp	GTT Val	TGG Trp	GAC Asp 425	TGG Trp	ATA Ile	TGC Cys	ACA Thr	GTT Val 430	TTG Leu	GCT Ala	1296
GAC Asp	TTC Phe	AAG Lys 435	ACC Thr	TGG Trp	CTC Leu	CAG Gln	TCC Ser 440	AAG Lys	CTC Leu	CTG Leu	CCG Pro	CGA Arg 445	TTA Leu	CCG Pro	GGA Gly	1344
GTC Val 450	CCC Pro	TTT Phe	TTC Phe	TCA Ser	TGC Cys	CAA Gln 455	CGT Arg	GGG Gly	TAC Tyr	AAG Lys	GGG Gly 460	GTC Val	TGG Trp	CGG Arg	GGA Gly	1392

GAC Asp 465	GGC Gly	ATC Ile	ATG Met	CAG Gln	ACC Thr 470	AGC Cys	TCA Ser	TGT Cys	GGA Gly 475	GCA Ala	CAG Gln	ACC Ile 480	GGA Gly	1440		
CAT His	GTC Val	AAA Lys	AAC Asn	GGT Gly 485	TCC Ser	ATG Met	AGG Arg	ATC Ile	GTT Val 490	GGG Gly	CCT Pro	AAG Lys	ACC Thr 495	TGT Cys	AGT Ser	1488
AAC Asn	ATG Met	TGG Trp	CAT His 500	GGA Gly	ACA Thr	TTC Phe	CCC Pro	ATC Ile 505	AAC Asn	GCA Ala	TAC Tyr	ACC Thr 510	ACG Thr	GGC Gly	CCC Pro	1536
TGC Cys	ACG Thr	CCC Pro 515	TCC Ser	CCA Pro	GCG Ala	CCA Pro 520	AAC Asn	TAT Tyr	TCC Ser	AGG Arg	GCG Ala	CTG Leu 525	TGG Trp	CGG Arg	GTG Val	1584
GCT Ala 530	GCT Ala	GAG Glu	GAG Glu	TAC Tyr	GTG Val	GAG Glu 535	GTT Val	ACG Thr	CGG Arg	GTG Val	GGG Gly 540	GAT Asp	TTC Phe	CAC His	TAC Tyr	1632
GTG Val 545	ACG Thr	AGC Ser	ATG Met	ACC Thr	ACT Thr 550	GAC Asp	AAC Asn	GTA Val	AAA Lys	TGC Cys 555	CCG Pro	TGC Cys	CAG Gln	GTT Val	CCA Pro 560	1680
GCC Ala	CCC Pro	GAA Glu	TTC Phe	TTC Phe 565	ACA Thr	GAA Glu	GTG Val	GAT Asp	GGG Gly 570	GTG Val	CGG Arg	CTG Leu	CAC His	AGG Arg 575	TAC Tyr	1728
GCT Ala	CCG Pro	GCG Ala	TGC Cys 580	AAA Lys	CCT Pro	CTC Leu	CTA Leu	CGG Arg 585	GAG Glu	GAG Glu	GTC Val	ACA Thr	TTC Phe 590	CAG Gln	GTC Val	1776
GGG Gly	CTC Leu	AAC Asn 595	CAA Gln	TAC Tyr	CTG Leu	GTT Val	GGG Gly 600	TCG Ser	CAG Gln	CTC Leu	CCA Pro	TGC Cys 605	GAG Glu	CCC Pro	GAA Glu	1824
CCG Pro	GAT Asp 610	GTA Val	GCA Ala	GTG Val	CTC Leu	ACT Thr 615	TCC Ser	ATG Met	CTC Leu	ACC Thr	GAC Asp 620	CCC Pro	TCC Ser	CAC His	ATC Ile	1872
ACA Thr 625	GCA Ala	GAG Glu	ACG Thr	GCT Ala	AAG Lys 630	CGC Arg	AGG Arg	CTG Leu	GCC Ala	AGG Arg 635	GGG Gly	TCT Ser	CCC Pro	CCC Pro	TCC Ser 640	1920
TTG Leu	GCC Ala	AGC Ser	TCT Ser	TCA Ser 645	GCT Ala	AGC Ser	CAG Gln	TTG Leu	TCT Ser 650	GCG Ala	CCT Pro	TCC Ser	TCG Ser	AAG Lys 655	GCG Ala	1968
ACA Thr	TAC Tyr	ATT Ile	ACC Thr 660	CAA Gln	AAT Asn	GAC Asp	TTC Phe	CCA Pro 665	GAC Asp	GCT Ala	GAC Asp	CTC Leu	ATC Ile	GAG Glu	GCC Ala 670	2016

AAC CTC CTG TGG CGG CAT G ATG GGC
 Asn Leu Leu Trp Arg His G Met Gly
 675 680

2043

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2116 base pairs
- (B) TYPE: nucleotide with corresponding protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to genomic RNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: cDNA clones from 5' end of the genome

(ix) FEATURE:

- (B) LOCATION: from 308 to 2116 bp start of the PT-NANBH polyprotein
- (D) OTHER INFORMATION: viral structural and non-structural proteins

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GATCACTCCC	CTGTGAGGAA	CTACTGTCTT	CACGCAGAAA	GCGTCTAGCC	ATGGCGTTAG	60		
TATGAGTGTC	GTGCAGCCTC	CAGGACCCCC	CCTCCCGGGA	GAGCCATAGT	GGTCTGCGGA	120		
ACCGGTGAGT	ACACCGGAAT	TGCCAGGACG	ACCGGGTCCT	TTCTTGGATT	AACCCGCTCA	180		
ATGCCTGGAG	ATTTGGGCGT	GCCCCGCAA	GACTGCTAGC	CGAGTAGTGT	TGGGTCGCGA	240		
AAGGCCTTGT	GGTACTGCCT	GATAGGGTGC	TTGCGAGTGC	CCCGGGAGGT	CTCGTAGACC	300		
GTGCACC	ATG AGC	ACG AAT	CCT AAA	CCT CAA	AGA AAA	ACC AAA	CGT AAC	349
	Met Ser	Thr Asn	Pro Lys	Pro Gln	Arg Lys	Thr Lys	Arg Asn	
			5		10			
ACC AAC	CGC CGC	CCA CAG	GAC GTC	AAG TTC	CCG GGC	GGT GGT	CAG ATC	397
Thr Asn	Pro Arg	Pro Gln	Asp Val	Lys Phe	Pro Gly	Gly Gly	Gln Ile	
15		20		25			30	
GTT GGT	GGA GTT	TAC CTG	TTG CCG	CGC AGG	GGC CCC	AGG TTG	GGT GTG	445
Val Gly	Gly Val	Tyr Leu	Leu Pro	Arg Arg	Gly Pro	Arg Leu	Gly Val	
		35		40			45	
CGC GCG	ACT AGG	AAG ACT	TCC GAG	CGG TCG	CAA CCT	CGT GGA	AGG CGA	493
Arg Ala	Thr Arg	Lys Thr	Ser Glu	Arg Ser	Gln Pro	Arg Gly	Arg Arg	
		50		55			60	

CAA Gln	CCT Pro	ATC Ile	CCC Pro	AAG Lys	GCT Ala	CGC Arg	AGG Gln	CCC Pro	GAG Glu	GGC Gly	AGG Arg	GCC Ala	TCT Trp	ACT Ala	GAG Gln	541
	65						70					75				
CCC Pro	GGG Gly	TAC Tyr	CCT Pro	TGG Trp	CCC Pro	CTC Leu	TAT Tyr	GGC Gly	AAC Asn	GAG Glu	GGC Gly	ATG Met	GGG Gly	TGG Trp	GCA Ala	589
	80						85					90				
GGA Gly	TGG Trp	CTC Leu	CTG Leu	TCA Ser	CCC Pro	CGT Arg	GGC Gly	TCC Ser	CGG Arg	CCT Pro	AGT Ser	TGG Trp	GGC Gly	CCC Pro	ACT Thr	637
100					105					110					115	
GAC Asp	CCC Pro	CGG Arg	CGT Arg	AGG Arg	TCG Ser	CGT Arg	AAT Asn	TTG Leu	GGT Gly	AAA Lys	GTC Val	ATC Ile	GAT Asp	ACC Thr	CTC Leu	685
				120					125					130		
ACA Thr	TGC Cys	GGC Gly	TTC Phe	GCC Ala	GAC Asp	CTC Leu	ATG Met	GGG Gly	TAC Tyr	ATT Ile	CCG Pro	CTC Leu	GTC Val	GGC Gly	GCT Ala	733
			135					140					145			
CCC Pro	TTA Leu	GGG Gly	GGC Gly	GCT Ala	GCC Ala	AGG Arg	GCC Ala	CTG Leu	GCG Ala	CAT His	GGC Gly	GTC Val	CGG Arg	GTT Val	CTG Leu	781
		150					155					160				
GAG Glu	GAC Asp	GGC Gly	GTG Val	AAC Asn	TAT Tyr	GCA Ala	ACA Thr	GGG Gly	AAT Asn	TTA Leu	CCC Pro	GGT Gly	TGC Cys	TCT Ser	TTC Phe	829
	165					170					175					
TCT Ser	ATC Ile	TTC Phe	CTC Leu	TTG Leu	GCT Ala	TTG Leu	CTG Leu	TCC Ser	TGT Cys	TTG Leu	ACC Thr	ATT Ile	CCA Pro	GCT Ala	TCC Ser	877
180					185					190					195	
GCT Ala	TAT Tyr	GAA Glu	GTG Val	CGC Arg	AAC Asn	GTG Val	TCC Ser	GGG Gly	ATC Ile	TAC Tyr	CAT His	GTC Val	ACG Thr	AAC Asn	GAT Asp	925
				200				205						210		
TGC Cys	TCC Ser	AAC Asn	TCA Ser	AGC Ser	ATC Ile	GTG Val	TAC Tyr	GAG Glu	ACA Thr	GCG Ala	GAC Asp	ATG Met	ATC Ile	ATG Met	CAC His	973
			215					220					225			
ACC Thr	CCC Pro	GGG Gly	TGT Cys	GTG Val	CCC Pro	TGT Cys	GTC Val	CGG Arg	GAG Glu	GGT Gly	AAT Asn	TCC Ser	TCC Ser	CGC Arg	TGC Cys	1021
		230					235					240				
TGG Trp	GTA Val	GCG Ala	CTC Leu	ACT Thr	CCC Pro	ACG Thr	CTC Leu	GCG Ala	GCC Ala	AAG Lys	GAC Asp	GCC Ala	AGC Ser	ATC Ile	CCC Pro	1069
	245					250					255					
ACT Thr	GCG Ala	ACA Thr	ATA Ile	CGA Arg	CGC Arg	CAC His	GTC Val	GAT Asp	TTG Leu	CTC Leu	GTT Val	GGG Gly	GCG Ala	GCT Ala	GCC Ala	1117
260					265					270					275	

TTC	TGC	TCC	GCT	ATG	TAC	GTT	TGG	GAT	CTC	TGC	GGA	TCT	GTT	TTC	CTC	1165
Phe	Cys	Ser	Ala	Met	Tyr	Val	Gly	Asp	Leu	Cys	Gly	Ser	Val	Phe	Leu	
			280						285					290		
GTC	TCT	CAG	CTG	TTC	ACC	TTC	TCG	CCT	CGC	CGA	CAT	CAG	ACG	GTA	CAG	1213
Val	Ser	Gln	Leu	Phe	Thr	Phe	Ser	Pro	Arg	Arg	His	Gln	Thr	Val	Gln	
			295					300					305			
GAC	TGC	AAT	TGT	TCA	ATC	TAT	CCC	GGC	CAC	GTA	TCA	GGT	CAC	CGC	ATG	1261
Asp	Cys	Asn	Cys	Ser	Ile	Tyr	Pro	Gly	His	Val	Ser	Gly	His	Arg	Met	
		310					315					320				
GCT	TGG	GAT	ATG	ATG	ATG	AAC	TGG	TCA	CCT	ACA	GCA	GCC	CTA	GTG	GTA	1309
Ala	Trp	Asp	Met	Met	Met	Asn	Trp	Ser	Pro	Thr	Ala	Ala	Leu	Val	Val	
	325					330					335					
TCG	CAG	CTA	CTC	CGG	ATC	CCA	CAA	GCT	GTC	GTG	GAC	ATG	GTG	GCG	GGG	1357
Ser	Gln	Leu	Leu	Arg	Ile	Pro	Gln	Ala	Val	Val	Asp	Met	Val	Ala	Gly	
340					345				350						355	
GCC	CAC	TGG	GGA	GTC	CTG	GCG	GGC	CTT	GCC	TAC	TAT	TCC	ATG	GTG	GGG	1405
Ala	His	Trp	Gly	Val	Leu	Ala	Gly	Leu	Ala	Tyr	Tyr	Ser	Met	Val	Gly	
				360					365					370		
AAC	TGG	GCT	AAG	GTC	TTG	GTT	GTG	ATG	CTA	CTC	TTT	GCC	GGC	GTT	GAC	1453
Asn	Trp	Ala	Lys	Val	Leu	Val	Val	Met	Leu	Leu	Phe	Ala	Gly	Val	Asp	
			375					380					385			
GGG	GAA	CCT	TAC	ACG	ACA	GGG	GGG	ACA	CAC	GGC	CGC	GCC	GCC	CAC	GGG	1501
Gly	Glu	Pro	Tyr	Thr	Thr	Gly	Gly	Thr	His	Gly	Arg	Ala	Ala	His	Gly	
		390					395					400				
CTT	ACA	TCC	CTC	TTC	ACA	CCT	GGG	CCG	GCT	CAG	AAA	ATC	CAG	CTT	GTA	1549
Leu	Thr	Ser	Leu	Phe	Thr	Pro	Gly	Pro	Ala	Gln	Lys	Ile	Gln	Leu	Val	
	405					410					415					
AAC	ACC	AAC	GGC	AGC	TGG	CAC	ATC	AAC	AGA	ACT	GCC	TTG	AAC	TGC	AAT	1597
Asn	Thr	Asn	Gly	Ser	Trp	His	Ile	Asn	Arg	Thr	Ala	Leu	Asn	Cys	Asn	
420					425					430					435	
GAC	TCC	CTC	CAA	ACT	GGG	TTC	CTT	GCC	GCG	CTG	TTC	TAC	ACG	CAC	AGG	1645
Asp	Ser	Leu	Gln	Thr	Gly	Phe	Leu	Ala	Ala	Leu	Phe	Tyr	Thr	His	Arg	
				440					445					450		
TTC	AAT	GCG	TCC	GGA	TGC	TCA	GAG	CGC	ATG	GCC	AGC	TGC	CGC	CCC	ATT	1693
Phe	Asn	Ala	Ser	Gly	Cys	Ser	Glu	Arg	Met	Ala	Ser	Cys	Arg	Pro	Ile	
			455					460					465			
GAC	CAG	TTC	GAT	CAG	GGG	TGG	GGT	CCC	ATC	ACT	TAT	AAT	GAG	TCC	CAC	1741
Asp	Gln	Phe	Asp	Gln	Gly	Trp	Gly	Pro	Ile	Thr	Tyr	Asn	Glu	Ser	His	
		470					475					480				

GGC	TTG	GAC	CAG	AGG	CCC	TAC	TGC	TGG	CAC	TAC	GCA	CCT	CCG	TGT	1789
Gly	Leu	Asp	Gln	Arg	Pro	Tyr	Cys	Trp	His	Tyr	Ala	Pro	Gln	Pro	Cys
485						490					495				
GGT	ATC	GTG	CCC	GCG	TTG	CAG	GTG	TGT	GGC	CCA	GTG	TAC	TGT	TTC	ACT
Gly	Ile	Val	Pro	Ala	Leu	Gln	Val	Cys	Gly	Pro	Val	Tyr	Cys	Phe	Thr
500					505					510					515
CCA	AGC	CCT	GTT	GTG	GTG	GGG	ACG	ACC	GAT	CGT	TTC	GGC	GCC	CCT	ACG
Pro	Ser	Pro	Val	Val	Val	Gly	Thr	Thr	Asp	Arg	Phe	Gly	Ala	Pro	Thr
				520					525					530	
TAC	AGA	TGG	GGT	GAG	AAT	GAG	ACG	GAC	GTG	CTG	CTT	CTC	AAC	AAC	ACG
Tyr	Arg	Trp	Gly	Glu	Asn	Glu	Thr	Asp	Val	Leu	Leu	Leu	Asn	Asn	Thr
			535					540					545		
CGG	CCG	CCA	CGG	GGC	AAC	TGG	TTC	GGC	TGT	ACA	TGG	ATG	AAT	AGC	ACC
Arg	Pro	Pro	Arg	Gly	Asn	Trp	Phe	Gly	Cys	Thr	Trp	Met	Asn	Ser	Thr
		550					555					560			
GGG	TTC	ACC	AAG	ACG	TGT	GGG	GGC	CCC	CCG	TGC	AAC	ATC	GGG	GGG	GTC
Gly	Phe	Thr	Lys	Thr	Cys	Gly	Gly	Pro	Pro	Cys	Asn	Ile	Gly	Gly	Val
565						570					575				
GGC	AAC	AAC	ACT	TTG	ATC	TGC	CCC	ACG	GAC	TGC	TTC	CGG	AAG	CAT	CCC
Gly	Asn	Asn	Thr	Leu	Ile	Cys	Pro	Thr	Asp	Cys	Phe	Arg	Lys	His	Pro
580					585					590					595
GAG	GCC	ACT	TAC	ACC	AAA	TGC	GGT	TCG	GGG	CCT	TGG	TTG			
Glu	Ala	Thr	Tyr	Thr	Lys	Cys	Gly	Ser	Gly	Pro	Trp	Leu			
				600					605						

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3750 base pairs
- (B) TYPE: nucleotide with corresponding protein
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to genomic RNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: human; serum infectious for PT-NANBH

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: cDNA clones from 3' end of the genome

(ix) FEATURE:

- (B) LOCATION: from 1 to 3750 bp portion of the PT-NANBH polyprotein
- (D) OTHER INFORMATION: viral non-structural proteins

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TGG Trp	GAG Glu	GGC Gly	GTC Val	TTC Phe	ACA Thr	GGC Gly	CTC Leu	ACC Thr	CAC His	GTG Val	GAT Asp	GCC Ala	CAC His	TTC Phe	CTG Leu	48
				5				10						15		
TCC Ser	CAA Gln	ACA Thr	AAG Lys	CAG Gln	GCA Ala	GGA Gly	GAC Asp	AAC Asn	TTC Phe	CCC Pro	TAC Tyr	CTG Leu	GTG Val	GCG Ala	TAC Tyr	96
			20					25					30			
CAG Gln	GCT Ala	ACT Thr	GTG Val	TGC Cys	GCT Ala	AGG Arg	GCC Ala	CAG Gln	GCC Ala	CCA Pro	CCT Pro	CCA Pro	TCA Ser	TGG Trp	GAT Asp	144
			35				40					45				
CAA Gln	ATG Met	TGG Trp	AAG Lys	TGT Cys	CTC Leu	ATA Ile	CGG Arg	CTA Leu	AAG Lys	CCT Pro	ACT Thr	CTG Leu	CGC Arg	GGG Gly	CCA Pro	192
	50					55					60					
ACA Thr	CCC Pro	TTG Leu	CTG Leu	TAT Tyr	AGG Arg	CTG Leu	GGA Gly	GCC Ala	GTC Val	CAA Gln	AAC Asn	GAG Glu	GTC Val	ACC Thr	CTC Leu	240
	65				70					75					80	
ACA Thr	CAC His	CCC Pro	ATA Ile	ACC Thr	AAA Lys	TTC Phe	ATC Ile	ATG Met	GCA Ala	TGC Cys	ATG Met	TCA Ser	GCC Ala	GAC Asp	CTG Leu	288
				85				90						95		
GAG Glu	GTC Val	GTC Val	ACG Thr	AGC Ser	ACC Thr	TGG Trp	GTG Val	CTG Leu	GTG Val	GGC Gly	GGG Gly	GTC Val	CTT Leu	GCA Ala	GCT Ala	336
			100					105					110			
CTG Leu	GCT Ala	GCG Ala	TAT Tyr	TGC Cys	TTG Leu	ACA Thr	ACA Thr	GGC Gly	AGC Ser	GTG Val	GTC Val	ATT Ile	GTG Val	GGT Gly	AGG Arg	384
			115				120					125				
ATC Ile	ATC Ile	TTG Leu	TCC Ser	GGG Gly	CGG Arg	CCG Pro	GCT Ala	ATT Ile	GTT Val	CCC Pro	GAC Asp	AGG Arg	GAA Glu	GTC Val	CTC Leu	432
	130					135					140					
TAC Tyr	CAG Gln	GAG Glu	TTC Phe	GAT Asp	GAG Glu	ATG Met	GAA Glu	GAG Glu	TGC Cys	GCG Ala	TCG Ser	CAC His	CTC Leu	CCT Pro	TAC Tyr	480
	145				150					155					160	
ATC Ile	GAG Glu	CAG Gln	GGA Gly	ATG Met	CAG Gln	CTC Leu	GCC Ala	GAG Glu	CAG Gln	TTC Phe	AAG Lys	CAA Gln	AAA Lys	GCG Ala	CTC Leu	528
			165					170						175		
GGG Gly	TTG Leu	CTG Leu	CAG Gln	ACA Thr	GCC Ala	ACC Thr	AAG Lys	CAA Gln	GCG Ala	GAG Glu	GCC Ala	GCT Ala	GCT Ala	CCC Pro	GTG Val	576
			180					185					190			
GTG Val	GAG Glu	TCC Ser	AAG Lys	TGG Trp	CGA Arg	GCC Ala	CTT Leu	GAG Glu	ACC Thr	TTC Phe	TGG Trp	GCG Ala	AAA Lys	CAC His	ATG Met	624
		195					200					205				

TGG Trp 210	AAC Asn 210	TTC Phe 210	ATC Ile 210	AGC Ser 210	GGG Gly 210	ATG Ile 215	CAG Gln 215	TAC Tyr 215	TTA Leu 215	GCA Ala 220	GGC Gly 220	TTG Leu 220	TCG Ser 220	ACT Thr 220	CTG Leu 220	672
CCT Pro 225	GGG Gly 225	AAT Asn 225	CCC Pro 225	GCG Ala 230	ATT Ile 230	GCA Ala 230	TCA Ser 230	CTG Leu 230	ATG Met 235	GCG Ala 235	TTC Phe 235	ACA Thr 235	GCC Ala 235	TCT Ser 240	GTC Val 240	720
ACT Thr 245	AGC Ser 245	CCG Pro 245	CTC Leu 245	ACC Thr 245	ACC Thr 245	CAA Gln 245	TCT Ser 250	ACC Thr 250	CTC Leu 250	CTG Leu 250	CTT Leu 250	AAC Asn 255	ATC Ile 255	CTG Leu 255	GGG Gly 255	768
GGA Gly 260	TGG Trp 260	GTA Val 260	GCC Ala 260	GCC Ala 260	CAA Gln 260	CTC Leu 265	GCT Ala 265	CCC Pro 265	CCC Pro 265	AGT Ser 265	GCT Ala 270	GCT Ala 270	TCA Ser 270	GCT Ala 270	TTC Phe 270	816
GTA Val 275	GGC Gly 275	GCC Ala 275	GGC Gly 275	ATT Ile 280	GCT Ala 280	GGT Gly 280	GCG Ala 280	GCT Ala 280	GTT Val 280	GGC Gly 285	AGC Ser 285	ATA Ile 285	GGC Gly 285	CTT Leu 285	GGG Gly 285	864
AAG Lys 290	GTG Val 290	CTT Leu 290	GTG Val 290	GAC Asp 295	ATC Ile 295	TTG Leu 295	GCG Ala 295	GGC Gly 295	TAT Tyr 300	GGA Gly 300	GCA Ala 300	GGA Gly 300	GTG Val 300	GCA Ala 300	GGC Gly 300	912
GCG Ala 305	CTC Leu 305	GTG Val 305	GCC Ala 305	TTT Phe 310	AAG Lys 310	GTC Val 310	ATG Met 310	AGC Ser 310	GGC Gly 315	GAA Glu 315	ATG Met 315	CCC Pro 315	TCC Ser 315	ACC Thr 320	GAG Glu 320	960
GAC Asp 325	CTG Leu 325	GTT Val 325	AAC Asn 325	TTA Leu 325	CTC Leu 325	CCT Pro 330	GCC Ala 330	ATC Ile 330	CTC Leu 330	TCT Ser 330	CCT Pro 330	GGT Gly 335	GCC Ala 335	CTG Leu 335	GTC Val 335	1008
GTC Val 340	GGG Gly 340	GTC Val 340	GTG Val 340	TGC Cys 340	GCA Ala 340	GCG Ala 345	ATA Ile 345	CTG Leu 345	CGT Arg 345	CGG Arg 345	CAC His 350	GTG Val 350	GGT Gly 350	CCA Pro 350	GGG Gly 350	1056
GAG Glu 355	GGG Gly 355	GCT Ala 355	GTG Val 355	CAG Gln 360	TGG Trp 360	ATG Met 360	AAC Asn 360	CGG Arg 360	CTG Leu 360	ATA Ile 365	GCG Ala 365	TTC Phe 365	GCC Ala 365	TCG Ser 365	CGG Arg 365	1104
GGT Gly 370	AAC Asn 370	CAT His 370	GTT Val 370	TCC Ser 375	CCC Pro 375	ACG Thr 375	CAC His 375	TAT Tyr 375	GTG Val 380	CCA Pro 380	GAG Glu 380	AGC Ser 380	GAC Asp 380	GCC Ala 380	GCA Ala 380	1152
GCA Ala 385	CGT Arg 385	GTC Val 385	ACT Thr 390	CAG Gln 390	ATC Ile 390	CTC Leu 390	TCC Ser 395	GAC Asp 395	CTT Leu 395	ACT Thr 395	ATC Ile 395	ACC Thr 400	CAA Gln 400	CTG Leu 400	TTG Leu 400	1200
AAG Lys 405	AGG Arg 405	CTC Leu 405	CAC His 405	CAG Gln 405	TGG Trp 410	ATT Ile 410	AAC Asn 410	GAG Glu 410	GAC Asp 410	TGC Cys 415	TCC Ser 415	ACG Thr 415	CCC Pro 415	TGC Cys 415	TCC Ser 415	1248

GGC	TCG	TGG	CTA	AGG	GAT	GTT	GGG	GAC	TGG	ATA	TGC	ACA	GTT	TTG	GCT	1296
Gly	Ser	Trp	Leu	Arg	Asp	Val	Trp	Asp	Trp	Ile	Cys	Thr	Val	Leu	Ala	
			420					425					430			
GAC	TTC	AAG	ACC	TGG	CTC	CAG	TCC	AAG	CTC	CTG	CCG	CGA	TTA	CCG	GGA	1344
Asp	Phe	Lys	Thr	Trp	Leu	Gln	Ser	Lys	Leu	Leu	Pro	Arg	Leu	Pro	Gly	
		435					440					445				
GTC	CCC	TTT	TTC	TCA	TGC	CAA	CGT	GGG	TAC	AAG	GGG	GTC	TGG	CGG	GGA	1392
Val	Pro	Phe	Phe	Ser	Cys	Gln	Arg	Gly	Tyr	Lys	Gly	Val	Trp	Arg	Gly	
	450					455					460					
GAC	GGC	ATC	ATG	CAG	ACC	ACC	TGC	TCA	TGT	GGA	GCA	CAG	ATC	ACC	GGA	1440
Asp	Gly	Ile	Met	Gln	Thr	Thr	Cys	Ser	Cys	Gly	Ala	Gln	Ile	Thr	Gly	
					470					475					480	
CAT	GTC	AAA	AAC	GGT	TCC	ATG	AGG	ATC	GTT	GGG	CCT	AAG	ACC	TGT	AGT	1488
His	Val	Lys	Asn	Gly	Ser	Met	Arg	Ile	Val	Gly	Pro	Lys	Thr	Cys	Ser	
				485					490					495		
AAC	ATG	TGG	CAT	GGA	ACA	TTC	CCC	ATC	AAC	GCA	TAC	ACC	ACG	GGC	CCC	1536
Asn	Met	Trp	His	Gly	Thr	Phe	Pro	Ile	Asn	Ala	Tyr	Thr	Thr	Gly	Pro	
			500					505					510			
TGC	ACG	CCC	TCC	CCA	GCG	CCA	AAC	TAT	TCC	AGG	GCG	CTG	TGG	CGG	GTG	1584
Cys	Thr	Pro	Ser	Pro	Ala	Pro	Asn	Tyr	Ser	Arg	Ala	Leu	Trp	Arg	Val	
		515					520					525				
GCT	GCT	GAG	GAG	TAC	GTG	GAG	GTT	ACG	CGG	GTG	GGG	GAT	TTC	CAC	TAC	1632
Ala	Ala	Glu	Glu	Tyr	Val	Glu	Val	Thr	Arg	Val	Gly	Asp	Phe	His	Tyr	
	530					535					540					
GTG	ACG	AGC	ATG	ACC	ACT	GAC	AAC	GTA	AAA	TGC	CCG	TGC	CAG	GTT	CCA	1680
Val	Thr	Ser	Met	Thr	Thr	Asp	Asn	Val	Lys	Cys	Pro	Cys	Gln	Val	Pro	
					550					555					560	
GCC	CCC	GAA	TTC	TTC	ACA	GAA	GTG	GAT	GGG	GTG	CGG	CTG	CAC	AGG	TAC	1728
Ala	Pro	Glu	Phe	Phe	Thr	Glu	Val	Asp	Gly	Val	Arg	Leu	His	Arg	Tyr	
				565					570					575		
GCT	CCG	GCG	TGC	AAA	CCT	CTC	CTA	CGG	GAG	GAG	GTC	ACA	TTC	CAG	GTC	1776
Ala	Pro	Ala	Cys	Lys	Pro	Leu	Leu	Arg	Glu	Glu	Val	Thr	Phe	Gln	Val	
			580					585					590			
GGG	CTC	AAC	CAA	TAC	CTG	GTT	GGG	TCG	CAG	CTC	CCA	TGC	GAG	CCC	GAA	1824
Gly	Leu	Asn	Gln	Tyr	Leu	Val	Gly	Ser	Gln	Leu	Pro	Cys	Glu	Pro	Glu	
		595					600					605				
CCG	GAT	GTA	GCA	GTG	CTC	ACT	TCC	ATG	CTC	ACC	GAC	CCC	TCC	CAC	ATC	1872
Pro	Asp	Val	Ala	Val	Leu	Thr	Ser	Met	Leu	Thr	Asp	Pro	Ser	His	Ile	
	610					615					620					

ACA Thr 625	GCA Ala	GAG Glu	ACG Thr	GCT Ala	AAG Lys 630	CGG Arg	GGG Arg	CTG Leu	GCC Ala	AGG Arg 635	GGG Gly	TCT Ser	CCC Pro	TCC Ser 640	1920	
TTG Leu	GCC Ala	AGC Ser	TCT Ser	TCA Ser 645	GCT Ala	AGC Ser	CAG Gln	TTG Leu	TCT Ser 650	GCG Ala	CCT Pro	TCC Ser	TCG Ser	AAG Lys 655	GCG Ala	1968
ACA Thr	TAC Tyr	ATT Ile	ACC Thr 660	CAA Gln	AAT Asn	GAC Asp	TTC Phe	CCA Pro 665	GAC Asp	GCT Ala	GAC Asp	CTC Leu	ATC Ile 670	GAG Glu	GCC Ala	2016
AAC Asn	CTC Leu	CTG Leu 675	TGG Trp	CGG Arg	CAT His	GAG Glu	ATG Met 680	GGC Gly	GGG Gly	GAC Asp	ATT Ile	ACC Thr 685	CGC Arg	GTG Val	GAG Glu	2064
TCA Ser 690	GAG Glu	AAC Asn	AAG Lys	GTA Val	GTA Val	ATC Ile 695	CTG Leu	GAC Asp	TCT Ser	TTC Phe	GAC Asp 700	CCG Pro	CTC Leu	CGA Arg	GCG Ala	2112
GAG Glu 705	GAG Glu	GAT Asp	GAG Glu	CGG Arg	GAA Glu 710	GTG Val	TCC Ser	GTC Val	CCG Pro	GCG Ala 715	GAG Glu	ATC Ile	CTG Leu	CGG Arg	AAA Lys 720	2160
TCC Ser	AAG Lys	AAA Lys	TTC Phe 725	CCA Pro	CCA Pro	GCG Ala	ATG Met	CCC Pro	GCA Ala 730	TGG Trp	GCA Ala	CGC Arg	CCG Pro	GAT Asp 735	TAC Tyr	2208
AAC Asn	CCT Pro	CCG Pro	CTG Leu 740	CTG Leu	GAG Glu	TCC Ser	TGG Trp 745	AAG Lys	GCC Ala	CCG Pro	GAC Asp	TAC Tyr 750	GTC Val	CCT Pro	CCA Pro	2256
GTG Val	GTA Val 755	CAT His	GGG Gly	TGC Cys	CCA Pro	CTG Leu	CCA Pro 760	CCT Pro	ACT Thr	AAG Lys	ACC Thr	CCT Pro 765	CCT Pro	ATA Ile	CCA Pro	2304
CCT Pro 770	CCA Pro	CGG Arg	AGG Arg	AAG Lys	AGG Arg	ACA Thr 775	GTT Val	GTT Val	CTG Leu	ACA Thr	GAA Glu 780	TCC Ser	ACC Thr	GTG Val	TCT Ser	2352
TCT Ser 785	GCC Ala	CTG Leu	GCG Ala	GAG Glu	CTT Leu 790	GCC Ala	ACA Thr	AAG Lys	GCT Ala	TTC Phe 795	GGT Gly	AGC Ser	TCC Ser	GAA Glu 800	CCG Pro	2400
TCG Ser	GCC Ala	GTC Val	GAC Asp 805	AGC Ser	GGC Gly	ACG Thr	GCA Ala	ACC Thr	GCC Ala 810	CCT Pro	CCT Pro	GAC Asp	CAA Gln	CCC Pro 815	TCC Ser	2448
GAC Asp	GAC Asp	GGC Gly	GGG Gly 820	GCA Ala	GGA Gly	TCT Ser	GAC Asp	GTT Val 825	GAG Glu	TCG Ser	TAT Tyr	TCC Ser	TCC Ser	ATG Met 830	CCC Pro	2496

CCC Pro	CTT Leu	GAG Glu 835	GGG Gly	GAG Glu	CCG Pro	GGG Gly 840	TAC Asp 840	CCC Pro	GAT Asp	CTC Leu	AGC Ser	GAC Asp 845	GGG Gly 845	TCT Ser	TGG Trp	2544
TCT Ser	ACC Thr 850	GTG Val	AGT Ser	GAG Glu	GAG Glu	GCC Ala 855	GGT Gly	GAG Glu	GAC Asp	GTC Val	GTC Val 860	TGC Cys	TGC Cys	TCG Ser	ATG Met	2592
TCC Ser 865	TAC Tyr	ACA Thr	TGG Trp	ACA Thr	GGC Gly 870	GCT Ala	CTG Leu	ATC Ile	ACG Thr	CCA Pro 875	TGC Cys	GCT Ala	GCG Ala	GAG Glu	GAA Glu 880	2640
AGC Ser	AAG Lys	CTG Leu	CCC Pro	ATC Ile 885	AAC Asn	GCG Ala	TTG Leu	AGC Ser	AAC Asn 890	TCT Ser	TTG Leu	CTG Leu	CGT Arg	CAC His 895	CAC His	2688
AAC Asn	ATG Met	GTC Val	TAC Tyr 900	GCT Ala	ACC Thr	ACA Thr	TCC Ser	CGC Arg 905	AGC Ser	GCA Ala	AGC Ser	CAG Gln	CGG Arg 910	CAG Gln	AAG Lys	2736
AAG Lys	GTC Val	ACC Thr 915	TTT Phe	GAC Asp	AGA Arg	CTG Leu	CAA Gln 920	ATC Ile	CTG Leu	GAC Asp	GAT Asp	CAC His 925	TAC Tyr	CAG Gln	GAC Asp	2784
GTG Val 930	CTC Leu	AAG Lys	GAG Glu	ATG Met	AAG Lys	GCG Ala 935	AAG Lys	GCG Ala	TCC Ser	ACA Thr	GTT Val 940	AAG Lys	GCT Ala	AAG Lys	CTT Leu	2832
CTA Leu 945	TCA Ser	GTA Val	GAG Glu	GAA Glu	GCC Ala 950	TGC Cys	AAG Lys	CTG Leu	ACG Thr	CCC Pro 955	CCA Pro	CAT His	TCG Ser	GCC Ala 960	AAA Lys	2880
TCT Ser	AAA Lys	TTT Phe	GGC Gly	TAT Tyr 965	GGG Gly	GCA Ala	AAG Lys	GAC Asp	GTC Val 970	CGG Arg	AAC Asn	CTA Leu	TCC Ser	AGC Ser 975	AAG Lys	2928
GCC Ala	ATT Ile	AAC Asn	CAC His 980	ATC Ile	CGC Arg	TCC Ser	GTG Val	TGG Trp 985	GAG Glu	GAC Asp	TTG Leu	TTG Leu	GAA Glu 990	GAC Asp	ACT Thr	2976
GAA Glu	ACA Thr	CCA Pro 995	ATT Ile	GAC Asp	ACC Thr	ACC Thr 1000	ATC Ile	ATG Met	GCA Ala	AAA Lys	AAT Asn 1005	GAG Glu 1005	GTT Val	TTC Phe	TGC Cys	3024
GTC Val 1010	CAA Gln 1010	CCA Pro	GAG Glu	AGA Arg	GGA Gly	GGC Gly 1015	CGC Arg	AAG Lys	CCA Pro	GCT Ala	CGC Arg 1020	CTT Leu	ATC Ile	GTG Val	TTC Phe	3072
CCA Pro 1025	GAC Asp	TTG Leu	GGG Gly	GTC Val 1030	CGT Arg	GTG Val	TGC Cys	GAG Glu	AAA Lys 1035	ATG Met	GCC Ala	CTC Leu	TAT Tyr	GAC Asp	GTG Val 1040	3120

GTC TCC ACC CTC CCT CAG GCT TTG ATG GGC TCC TCG TAC GGT TTC CAG	3168
Val Ser Thr Leu Pro Gln Ala Val Met Gly Ser Ser Tyr Gln Phe Gln	
1045 1050 1055	
TAT TCT CCT GGA CAG CGG GTC GAG TTC CTG GTG AAC GCC TGG AAA TCA	3216
Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu Val Asn Ala Trp Lys Ser	
1060 1065 1070	
AAG AAG ACC CCT ATG GGC TTT GCA TAT GAC ACC CGC TGT TTT GAC TCA	3264
Lys Lys Thr Pro Met Gly Phe Ala Tyr Asp Thr Arg Cys Phe Asp Ser	
1075 1080 1085	
ACA GTC ACT GAG AAT GAC ATC CGT GTA GAG GAG TCA ATT TAT CAA TGT	3312
Thr Val Thr Glu Asn Asp Ile Arg Val Glu Glu Ser Ile Tyr Gln Cys	
1090 1095 1100	
TGT GAC TTG GCC CCC GAA GCC AGA CAG GCC ATA AGG TCG CTC ACA GAG	3360
Cys Asp Leu Ala Pro Glu Ala Arg Gln Ala Ile Arg Ser Leu Thr Glu	
1105 1110 1115 1120	
CGG CTT TAT ATC GGG GGT CCC CTG ACT AAT TCA AAA GGG CAG AAC TGC	3408
Arg Leu Tyr Ile Gly Gly Pro Leu Thr Asn Ser Lys Gly Gln Asn Cys	
1125 1130 1135	
GGC TAT CGC CGG TGC CGC GCG AGC GGC GTG CTG ACG ACT AGC TGC GGT	3456
Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Cys Gly	
1140 1145 1150	
AAT ACC CTC ACA TGT TAC TTG AAG GCC TCT GCA GCC TGT CGA GCT GCA	3504
Asn Thr Leu Thr Cys Tyr Leu Lys Ala Ser Ala Ala Cys Arg Ala Ala	
1155 1160 1165	
AAG CTC CAG GAC TGC ACG ATG CTC GTG TGC GGA GAC GGC CTT GTC GTT	3552
Lys Leu Gln Asp Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val	
1170 1175 1180	
ATC TGT GAG AGC GCG GGA ACC CAG GAG GAC GCG GCG AGC CTA CGA GTC	3600
Ile Cys Glu Ser Ala Gly Thr Gln Glu Asp Ala Ala Ser Leu Arg Val	
1185 1190 1195 1200	
TTC ACG GAG GCT ATG ACT AGG TAC TCT GCC CCC CCC GGG GAC CCG CCC	3648
Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro	
1205 1210 1215	
CAA CCA GAA TAC GAC CTG GAG TTG ATA ACA TCA TGC TCC TCC AAT GTG	3696
Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val	
1220 1225 1230	
TCG GTC GCG CAC GAT GCA TCT GGC AAA AGG GTA TAC TAC CTC ACC CGT	3744
Ser Val Ala His Asp Ala Ser Gly Lys Arg Val Tyr Tyr Leu Thr Arg	
1235 1240 1245	

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: baculovirus AcNPV

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: oligonucleotide synthesizer; oligo d24

(ix) FEATURE:

- (B) LOCATION: from 1 to 23 bases homologous to portion of AcNPV polyhedrin gene downstream of the BamH1 cloning site in pAc360 and similar vectors
- (D) OTHER INFORMATION: primes DNA synthesis from baculovirus transfer vector sequences which flank DNA inserted at the BamH1 site.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CGGGTTTAAC ATTACGGATT TCC

23

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: baculovirus AcNPV

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: oligonucleotide synthesizer; oligo d126

(ix) FEATURE:

(B) LOCATION: from 1 to 31 bases homologous to the upstream junction sequences produced when cDNA amplified by d75 (SEQ ID NO:5) is cloned into the BamH1 cloning site in pAc360 and similar vectors; mismatches at bases 13 and 14 introduce a Pst1 site from 1 to 10 bases.

(D) OTHER INFORMATION: primes DNA synthesis at the junction of baculovirus transfer vector sequences and sequences previously amplified by oligo d75; introduces a Pst1 recognition site for subsequent cloning work

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TAAGGATCCC CCT GCA GTA TCG GCG GAA TTC
Ser Ala Val Ser Ala Glu Phe
5

31

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: N/A

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: oligonucleotide synthesizer; oligo d132

(ix) FEATURE:

(B) LOCATION: form 5 to 10 bases Pst1 recognition site
from 13 to 27 bases linker coding for five Lys residues
from 28 to 45 bases homologous to bases 4 to 21 of BR11
(SEQ ID NO:7)

(D) OTHER INFORMATION: primes DNA synthesis at the 5' end of BR11 and introduces a synthetic sequence which codes for five lysines as well as a Pst1 recognition site for subsequent cloning work

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CTGCCTGCA GTA AAG AAG AAG AAG AAG AAA ACC AAA CGT AAC ACC A
Val Lys Lys Lys Lys Lys Lys Thr Lys Arg Asn Leu
5 10

46

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.